

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-047950

(43)Date of publication of application : 18.02.2000

(51)Int.Cl.

G06F 13/00
G06F 13/14

(21)Application number : 10-225445

(71)Applicant : HITACHI LTD

(22)Date of filing : 24.07.1998

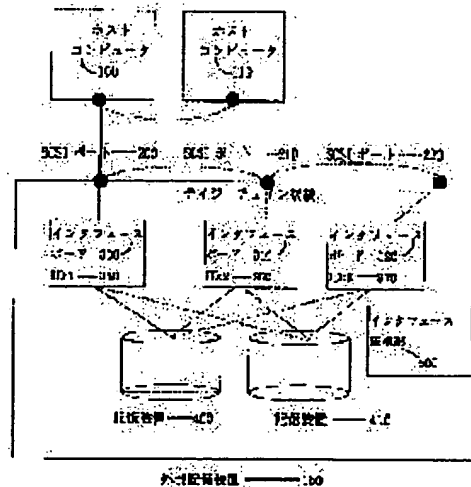
(72)Inventor : URATANI IKUO
URABE KIICHIRO

(54) EXTERNAL STORAGE DEVICE

(57)Abstract:

PROBLEM TO BE SOLVED: To eliminate the burden of switching to a host computer while keeping high reliability to the fault of plural paths with which storage devices are connected to each other.

SOLUTION: An input/output request and transfer data from a host computer 100 are transmitted/received through an SCSI port 200 and SCSI interface board 300. Storage devices 400-410 are connected to each other in an array shape so as to be able to be shared from plural SCSI interface boards 300-320. An interface monitoring device 500 monitors conditions of interface boards 300-320 at all times, and when any abnormality is found out, the interface board of light statistic load is selected out of normally operating interface boards so that the ID switching of interface boards 300-320 is performed between the both. Thus, the user of an external storage device 150 can continuously perform processing even in the case of operation defect on the interface board and the reliability of the device is improved as a whole.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-47950

(P2000-47950A)

(43) 公開日 平成12年2月18日 (2000.2.18)

| (51) Int.Cl. ⁷ | 識別記号 | F I | テ-マ-ト [*] (参考) |
|---------------------------|-------|---------------|-------------------------|
| G 0 6 F 13/00 | 3 0 1 | G 0 6 F 13/00 | 3 0 1 M 5 B 0 1 4 |
| 13/14 | 3 1 0 | 13/14 | 3 1 0 E 5 B 0 8 3 |

審査請求 未請求 請求項の数 3 F D (全 6 頁)

(21) 出願番号 特願平10-225445
(22) 出願日 平成10年7月24日 (1998.7.24)

(71) 出願人 000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目6番地
(72) 発明者 裏谷 郁夫
神奈川県小田原市国府津2880番地 株式会
社日立製作所ストレージシステム事業部内
(72) 発明者 占部 喜一郎
神奈川県小田原市国府津2880番地 株式会
社日立製作所ストレージシステム事業部内
(74) 代理人 100099302
弁理士 笹岡 茂 (外1名)

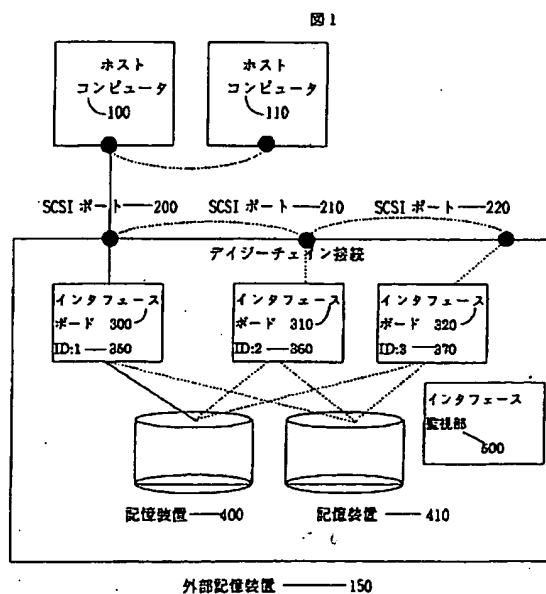
最終頁に続く

(54) 【発明の名称】 外部記憶装置

(57) 【要約】

【課題】 外部記憶装置について、記憶装置を接続する複数のパスの障害に対する高い信頼性を保持しつつ、ホストコンピュータに切り替えの負担を無くす。

【解決手段】 ホストコンピュータ100からの入出力要求および転送データはSCSIポート200を通じ、SCSIインタフェースボード300を介して送受信される。記憶装置(400-410)は、アレイ状に接続され、複数のSCSIインタフェースボード(300-320)から共有が可能な接続をしている。インタフェース監視装置500はインタフェースボード(300-320)の状況を常時監視し、異常が発見されたときは、正常動作をしているインタフェースボードの中から統計上負荷の軽いインタフェースボードを選択し、両者の間でインタフェースボード(300-320)のID切り替えを行う。これにより、外部記憶装置150使用者はインタフェースボードの動作不良の際にも継続して処理を行う事が可能となり、装置全体の信頼性が向上する。



1

2

【特許請求の範囲】

【請求項1】 上位装置との間で授受されるデータが格納される一以上の記憶装置と、該記憶装置と該上位装置との間に介在し該上位装置と該記憶装置との間における前記データの授受を制御する複数の記憶制御装置とを含む外部記憶装置であって、

前記各記憶制御装置には、前記上位装置が該記憶制御装置の指定に用いる識別子が各々設定されており、

複数の前記記憶装置が前記上位装置から見て等価に見えるように前記記憶制御装置を前記上位装置に接続するインタフェース手段と、

前記全ての記憶制御装置に接続され、前記記憶制御装置における障害の有無を監視し、障害有りのとき、障害有りの記憶制御装置と障害無しの一つの記憶制御装置との間で夫々が有する識別子を互いに交換設定する監視手段を備えることを特徴とする外部記憶装置。

【請求項2】 請求項1記載の外部記憶装置において、前記インタフェース手段は前記各記憶制御装置対応に設けられ、前記各記憶制御装置に対してそれぞれ別々の識別子を設定する手段を備えることを特徴とする外部記憶装置。

【請求項3】 請求項2記載の外部記憶装置において、前記記憶制御装置がそれぞれ接続された前記複数のインタフェース手段はデジチチェーン接続されていることを特徴とする外部記憶装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、コンピュータ装置または制御装置に接続される外部記憶装置に係わり、高い信頼性を持つ外部記憶装置に関する。

【0002】

【従来の技術】外部記憶装置の無停止稼働および無停止保守の実現という信頼性に対する要求から、インタフェースが故障した場合に、処理するインタフェースをホストコンピュータ側で切り替えて再アクセスすることにより処理を続行するというのが一般的である。この場合はオープンプラットフォームの場合、バスの切り替えが自動で実行できないため、故障検出、切り替え、データ修復に多くの時間を費やす必要がある。特開平8-335155号公報は、無停止稼働について述べられている。この方法では単一の識別子(ID)に対して冗長構成を用いることで対応することが可能である。しかし、この方法では複数のIDに対応するためにはID毎に冗長構成が必要であり、単一のIDに対する冗長構成が全て動作停止する際に対処することができなかった。

【0003】

【発明が解決しようとする課題】最近の市場動向として、高い信頼性を持つ外部記憶装置を接続した多目的ファイルサーバに対する要求が高まってきている。クライアントに対して複数のディスクイメージを見せることが

できることから、高い信頼性を伴った重要データに対するトランザクション処理への対応を要求される。そのためには、外部記憶装置内部の記憶装置に対して、高い信頼性の実現が必須である。しかし、従来の技術ではオープンプラットフォームの機器故障に対するバス切り替えの煩雑さに対して考慮されていなかった。本発明の課題は高い信頼性が要求される外部記憶装置について、記憶装置を接続する複数のバスの障害に対する高い信頼性を保持しつつ、ホストコンピュータに切り替えの負担をかけないことである。

【0004】

【課題を解決するための手段】上位装置との間で授受されるデータが格納される一以上の記憶装置と、該記憶装置と該上位装置との間に介在し該上位装置と該記憶装置との間における前記データの授受を制御する複数の記憶制御装置とを含む外部記憶装置であり、前記各記憶制御装置には、前記上位装置が該記憶制御装置の指定に用いる識別子が各々設定されており、複数の前記記憶装置が前記上位装置から見て等価に見えるように前記記憶制御装置を前記上位装置に接続するインタフェース手段と、前記全ての記憶制御装置に接続され、前記記憶制御装置における障害の有無を監視し、障害有りのとき、障害有りの記憶制御装置と障害無しの一つの記憶制御装置との間で夫々が有する識別子を互いに交換設定する監視手段を備えるようにしている。

【0005】また、前記インタフェース手段は前記各記憶制御装置対応に設けられ、前記各記憶制御装置に対してそれぞれ別々の識別子を設定する手段を備えるようにしている。また、前記記憶制御装置がそれぞれ接続された前記複数のインタフェース手段はデジチチェーン接続されているようにしている。

【0006】

【発明の実施の形態】以下、本発明に係る外部記憶装置の実施例について説明する。図1は本実施例に関わる外部記憶装置を含むシステム全体の構成を示した図である。外部記憶装置150は複数の記憶制御装置に相当するSCSIインタフェースボード(300-320)とそれぞれに直結するインタフェース手段に相当するSCSIポート(200-220)、SCSIインタフェースボードを監視するインタフェース監視部500、複数の記憶装置(400-410)で構成される。

【0007】複数の記憶装置(400-410)と複数のSCSIインタフェースボード(300-320)とSCSIポート(200-220)は自由に接続の構成が設定できる。この設定は、図示はしていないが、外部記憶装置内に配置された接続のためのスイッチを、外部記憶装置に設けたマイクロプロセッサ等により制御することにより行うことができ、また、外部記憶装置に設けたスイッチボードを手動により制御することにより行うこともできる。これにより、単一の記憶装置(400-410)に対する複数のインタフェースボード(300-320)からの共有が可能となっている。SCSI ID(350

-370)はインタフェースボード (300-320)毎に別々の値を持ち、SCSI ID(350-370)は、図示はしていないが、外部記憶装置に設けたマイクロプロセッサ等により設定される。インタフェースボード (300-320)の接続されたSCSIポート (200-220)はデジタイチェーンで接続されており、個々のインタフェースボード (300-320)が動作可能であるときはホストコンピュータ (100-110)から別々の処理を並行して実行する。

【0008】インタフェース監視部500は全てのインタフェースボード (300-320)と接続され、通信しており、
10 インタフェースボード (300-320)の内部状況を一定間隔で常に受け取り、解析する事により動作確認のチェックを行い、インタフェースボード (300-320)の動作不良が確認された場合には外部への連絡と、SCSI ID(350-370)の交換を動的に行うことができる。インタフェースボード動作時はそれぞれのSCSI IDへのI/O要求を指定されたインタフェースボードが受け取り、設定により接続された記憶装置に対して要求に対応する処理を行う。

【0009】図1の場合、ホストコンピュータ100によるSCSI ID 1への要求はインタフェースボード 300が受け
20 取り、記憶装置400に要求に対する処理を行う。しかし、故障もしくは保守のためにインタフェースボード 300が動作しなくなったときを考える。このとき、インタフェース監視部500はインタフェースボード 300からの内部状況の送信が行われず、もしくは受信した内部状況に異常を検出する事からインタフェースボード 300の動作不良を発見する。

【0010】インタフェース監視部500は、正常動作をしているインタフェースボード (310,320)の中から統計上負荷(ここで負荷とは、一定時間に動作した割合をいう)の軽いインタフェースボード (ここでは320)を選択し、インタフェースボード 320が動作不良でない事を確認した後、インタフェースボード 300と320のSCSI ID350と370を交換し、同時に、この交換に伴い、前記図示していないマイクロプロセッサ等により設定情報に基づきインタフェースボード 300および320と記憶装置 (400-410)との接続も交換する。この交換によりホストコンピュータ100はインタフェースボード 300の故障を意識する事なく処理の続行が可能となる。

【0011】SCSI IDの切換えはホストコンピュータのSCSIコマンド実行失敗によるSCSIコマンドの再送信の間
40 に実行する。外部記憶装置へのSCSIコマンドの再送信の間隔は数十秒単位であるため、それより十分に短い期間であればインタフェースボードを切り離してSCSIバス上にコマンド再送信先のSCSI IDが存在していなくとも影響はない。そのためインタフェースボード 300の障害のときには一旦インタフェースボード 300と320をSCSIバスから切り離し、ID設定の交換を終了させた後、再度SCSIバスに接続するという作業を行うことでインタフェースボードのID交換とする。

【0012】図2は、上記IDの交換および該交換に伴うインタフェースボードと記憶装置との接続の交換をした後の外部記憶装置を含むシステム全体の構成を示す。

【0013】SCSI ID 1と3の交換が既に行われた状態で、SCSI ID 3(SCSI ID 3の設定されたインタフェースボードは上記のとおり動作不良となっている)について要求が発生したときにも同様にインタフェースボード 310と320のうち負荷の軽いボードがID切り替えを行い処理を続行する。したがって、同様の処理をインタフェース
10 ボード動作不良が複数発生したときのことを考えると、外部記憶装置中に1つのインタフェースボードが正常動作可能であれば外部記憶装置として動作が可能となる。

【0014】図3はインタフェース監視部500の内部構成を示すブロック図である。マイクロプロセッサユニット510(以下、MPUと略す)は、ランダムアクセスメモリ520(以下、RAMと略す)に格納されているプログラム(情報を含む)を逐次読み出し、デコードしながら実行し、接続機構530を通じてインタフェースボード 300-320の状態監視およびID入れ替えの制御を実行している。接続機構530はMPUからの要求によりインタフェース
20 ボード 300-320と接続し、インタフェースボード 300-320の動作状況を受け取ることができる。

【0015】以下、図4のフローチャートにより前述の監視の手続きの一例を説明する。インタフェース監視部500は、デジタイチェーンされた全てのSCSIポート200-220へ上位装置からの入力が行われると、それを検出したインタフェースボード全てが接続機構530を介してMPU510に割込みを送り(ただし、割込みマスクにより割込みの多重は起こらない)、MPU510を起動し(ステップ600)、ステップ601で上位装置からの入力により指定されたID、すなわち外部要求が指定するID、に相当するインタフェースボードの状態を接続機構530より受け取る。ステップ602にて状態の内容に矛盾が無いかを判断し、インタフェースボードが動作している場合はステップ607にて受け取った負荷情報をRAMに格納する。

【0016】インタフェースボードが動作していない場合はステップ603で現在動作可能でかつ使用中または要求待ちのロック状態以外のインタフェースボードのうちもっとも負荷の軽いものをRAM内の負荷情報を参照して選び出す。ステップ605で選び出したボードのIDと動作していないボードとのIDの交換を行い、ステップ606で、ステップ605で選び出したボードが要求待ちのロック状態であることをRAMに格納する。また、前述の条件のポートが存在しない場合は終了し、外部からの要求の再送を待ち、その際に再度切換えを行う。ここで動作可能であるインタフェースボードが1つのみの場合に動作していないIDを指定した要求が交互に外部記憶装置100へ送られると、前述の交互のID間でID切換えが605で際限なく実行され、要求の実行が全く行われない事態が考えられる。使用される要求待ちロック状態はこれを避け
50

5

るために必要である。このロック状態は外部からの要求を1つ実行するまで変更しない。

【0017】

【発明の効果】本発明により、複数の記憶装置をアレイ状に接続した構成を持つ外部記憶装置について、装置使用者は装置を構成する記憶装置の実装状態を意識することなく、インタフェース異常の際にも継続して処理を行う事が可能となり、装置全体の信頼性が向上する。

【図面の簡単な説明】

【図1】外部記憶装置を含むシステム全体の構成を示す図である。

【図2】ID切り替え後のシステム全体の構成を示す図である。

【図3】インタフェース監視部の内部構成を示すブロック図である。

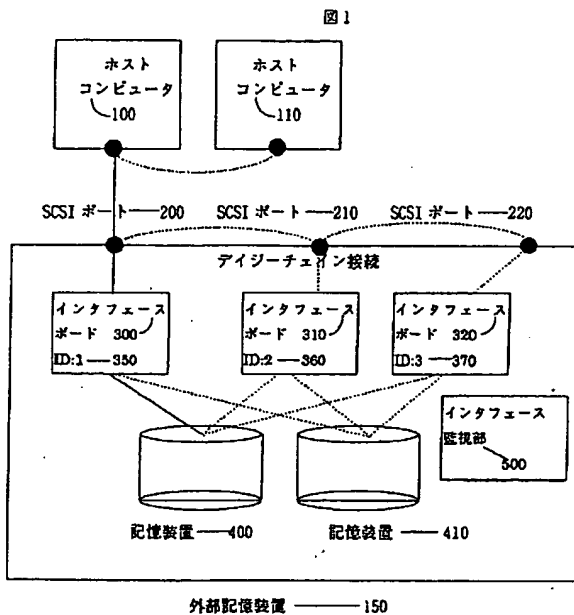
6

【図4】インタフェース監視部における監視の手続きのフローチャートを示す図である。

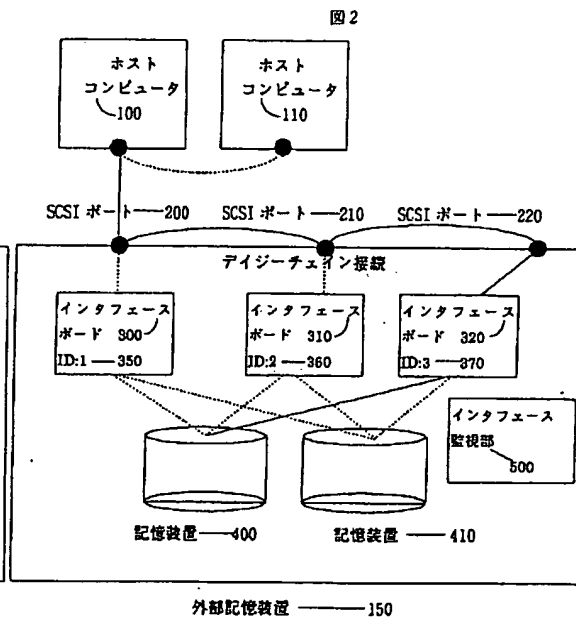
【符号の説明】

100～110 ホストコンピュータ
150 外部記憶装置
200～220 SCSIポート
300～320 インタフェースボード
350～370 インタフェースボードのSCSI ID
400～410 記憶装置
500 インタフェース監視部
510 マイクロプロセッサユニット(MPU)
520 ランダムアクセスメモリ(RAM)
530 接続機構

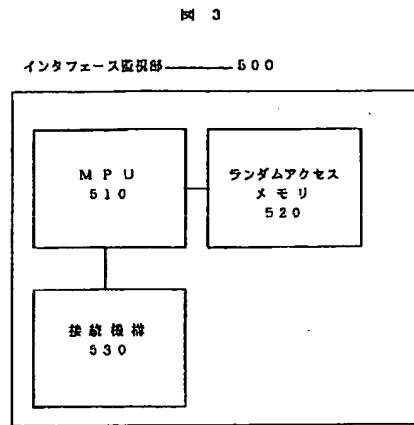
【図1】



【図2】

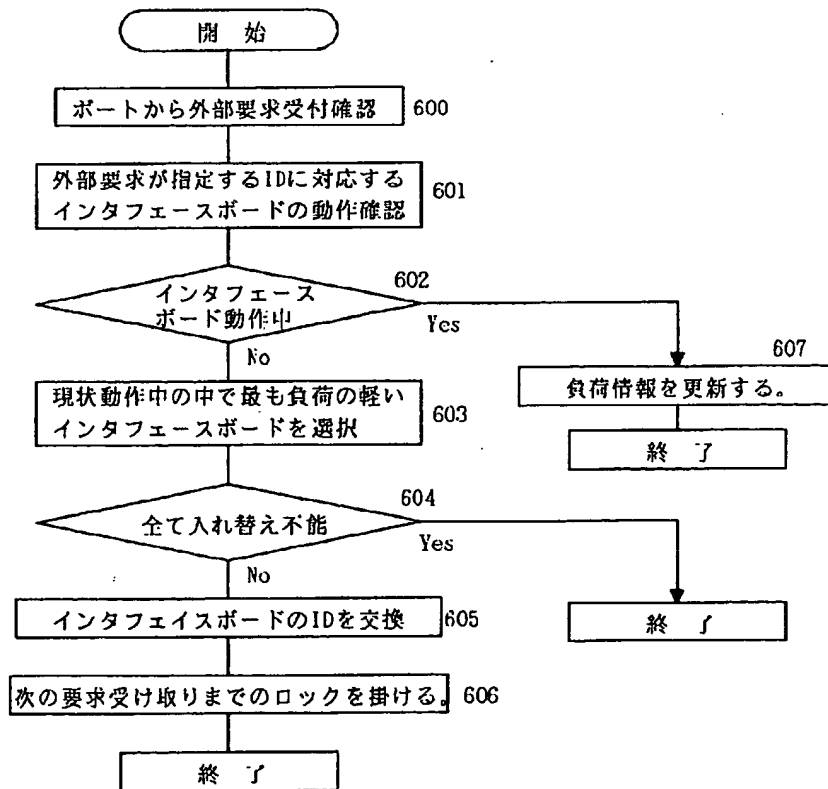


【 図3 】



【 図4 】

図 4



フロント ページの続き

F ターム (参考) 5B014 EA04 EB04 GD05 GD06 GD21
GD32 HA09 HB01 HB14 HB15
HC12 HC13
5B083 AA00 BB01 BB03 BB11 CD09
CD10 CD11 EE06 EE11 GG04